

Li, H., Shrestha, A., Fioranelli, F. , Le Kernec, J. , Heidari, H. , Pepa, M., Cippitelli, E., Gambi, E. and Spinsante, S. (2017) Multisensory Data Fusion for Human Activities Classification and Fall Detection. In: IEEE Sensors 2017, Glasgow, UK, 30 Oct - 01 Nov 2017, ISBN 9781509010127 (doi:[10.1109/ICSENS.2017.8234179](https://doi.org/10.1109/ICSENS.2017.8234179))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/148716/>

Deposited on: 08 February 2018

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Multisensor Data Fusion for Human Activities Classification and Fall Detection

Haobo Li, Aman Shrestha, Francesco Fioranelli, Julien Le Kernec, Hadi Heidari
School of Engineering, University of Glasgow,
Glasgow, United Kingdom

Matteo Pepa, Enea Cippitelli, Ennio Gambi, Susanna Spinsante
Department of Information Engineering, Università
Politecnica delle Marche, Ancona, Italy

Abstract—Significant research exists on the use of wearable sensors in the context of assisted living for activities recognition and fall detection, whereas radar sensors have been studied only recently in this domain. This paper approaches the performance limitation of using individual sensors, especially for classification of similar activities, by implementing information fusion of features extracted from experimental data collected by different sensors, namely a tri-axial accelerometer, a micro-Doppler radar, and a depth camera. Preliminary results confirm that combining information from heterogeneous sensors improves the overall performance of the system. The classification accuracy attained by means of this fusion approach improves by 11.2% compared to radar-only use, and by 16.9% compared to the accelerometer. Furthermore, adding features extracted from a RGB-D Kinect sensor, the overall classification accuracy increases up to 91.3%.

Keywords—Accelerometer, Radar sensor, Depth camera, human activities classification, fall detection, machine learning, data fusion.

I. INTRODUCTION

With increasingly aging population worldwide and increasing incidence of multi-morbidity conditions (i.e. the simultaneous presence of more than several chronic health issues), there is a significant need for automatic systems and sensors capable of classifying human activities and promptly detecting critical events such as falls [1]. Falls have obvious physical consequences, and there is a proven correlation between the long-lie time spent on the floor after the event and a reduction in life expectancy [2]. Activity classification can help characterise a normal behaviour pattern for the monitored patients, to detect anomalies that may be linked to deteriorating physical or cognitive health.

Different sensors have been proposed for the aforementioned applications in the context of Ambient Assisted Living (AAL) [3], namely wearables such as accelerometers, gyroscopes and magnetic sensors [4,5,6], video-camera sensors [5], and depth cameras and radar sensors [7] among others. To select one of the many possible technologies, one has to consider the different advantages and disadvantages of each type of sensor, in terms of performance (classification accuracy, rejection of false alarms, and percentage of missed detection), and regarding aspects such as end-users' acceptance, cost, power consumption, easiness of use and deployment. This leads to the investigation of how information from heterogeneous sensors can be used, leveraging on the strengths of each of them through information fusion. The rest of this paper presents

preliminary results of this investigation, with experimental data collected using a radar sensor, a RGB-D Kinect, and a tri-axial accelerometer within a smartphone. The main contribution is the initial investigation of fusion information from heterogeneous sensors, including radar sensors, as well as in the use of a rich set of experimental data, both in terms of the activities considered for classification and the number and age span of the participants (compared with other studies in the literature for radar and wearables [5]).

II. DATA COLLECTION AND FEATURE EXTRACTION

The radar sensor is an off-the-shelf Frequency Modulated Continuous Wave (FMCW) radar operating at 5.8 GHz and capable of recording micro-Doppler signatures of the targets of interest, i.e. Doppler vs time patterns of moving targets [8]. Microsoft Kinect sensor estimates the coordinates of joints corresponding to different body parts of the monitored subject and records their temporal evolution frame by frame. The tri-axial accelerometer within a commercial smartphone samples and records linear acceleration along the X, Y, and Z axis at approximately 100 Hz (maximum rate allowed by the smartphone). We recorded 10 different activities indicated in Table I, with 16 participants with age from 23 to 58 years. The measurements were collected in an office environment at Glasgow University. These activities were selected to be similar in pairs (e.g. 1 and 2, or 7 and 8) for an additional classification challenge, and to trigger possible false alarms when detecting falls, for example activity 3, 6, and 10, all presenting a fast acceleration component directed towards the floor.

TABLE I. LIST OF HUMAN ACTIVITIES

No	Description
1	Walking back and forth
2	Walking and carrying an object with both hands
3	Sitting down on a chair
4	Standing up from a chair
5	Bending to pick up an object and coming back up
6	Bending and staying down to tie shoelaces
7	Drinking a glass of water while standing
8	Picking up a phone call while standing
9	Simulating tripping and falling down frontally
10	Bending to check under furniture and coming back up

Numerical parameters, referred to as features, were extracted from the data of each sensor. For the tri-axial accelerometer these features were inspired from previous work in this domain [9-11] and are summarised in table II, divided into time and frequency domain. Frequency domain features aim to capture

the spectral energy distribution and include the amplitude of the Power Spectral Density (PSD) at a selected frequency band, the sum of the Fast Fourier Transform (FFT) coefficients, and the spectral entropy based on the power spectrum. For the radar sensor, the spectrograms (Doppler vs time patterns) were calculated by applying STFT (Short Time Fourier Transform) to the raw range-time radar data, and then features were extracted from the resulting images. These are summarised in table III. Entropy and skewness are related to the energy distribution of the pixels in the spectrogram. The centroid estimates the centre of mass of the micro-Doppler signature in the spectrogram, whereas the bandwidth estimates the energy content around it. The SVD (Singular Value Decomposition) of the spectrograms projects the energy content onto vectors in time and frequency domain. Statistical moments (mean and variance) of the first three left and right vectors are considered as features, as well as those of the centroid and bandwidth parameters over time.

TABLE II. TABLE OF FEATURES FOR ACCELEROMETER SENSOR

Time domain	No.	Frequency domain	No.
Mean	3	Spectral Power	9
Standard Deviation	3	Coefficients Sum	3
Autocorrelation	3	Spectral Entropy	3
Cross Correlation	3		
Variance	3		
RMS (Root Mean Square)	3		
MAD (Median Absolute Deviation)	3		
Inter-quadrature Range	3		
Range	3		
Minimum	3		
Number of features	30	Number of features	15

TABLE III. TABLE OF FEATURES FOR RADAR SENSOR

Radar Features	No.
Entropy of spectrogram	1
Skewness of spectrogram	1
Centroid of spectrogram (mean & variance)	2
Bandwidth of spectrogram (mean & variance)	2
Singular Value Decomposition (mean & variance of right and left vectors)	13
Number of features	19

III. RESULTS ANALYSIS USING RADAR AND ACCELEROMETER

The features listed in section II are used as input to classifiers based on supervised machine learning. A 16-fold cross-validation approach was applied, whereby data from 15

participants were used for training and data from the 16th remaining participant were used for testing. This was repeated 16 times, one for each participant, and the average classification results are reported in this paper.

Fig. 1(a)-1(c) shows confusion matrices to validate the classification performance of the proposed approach. The rows show the true classes of the 10 activities under test, whereas the columns show the predicted classes as estimated by the SVM quadratic-kernel classifier. An ideal result would have 100% classification on the diagonal of the confusion matrices, whereas the elements outside are related to misclassification events. A colour code is used to highlight in green the desired elements on the diagonal, and in yellow or in red colour the misclassifications. Activity 9 is highlighted as well, as this is the simulated fall activity. The radar sensor generates the majority of misclassifications for activity 1 and 2, but also activities 4, 5, and 8 contribute to several mistakes. The average classification is 68.8%. For the accelerometer, activity 3, 5, 6, and 8 are the most problematic ones for classification, and the average classification accuracy is 63.1% across the ten activities. For both sensors, the classification accuracy for activity 9 (the fall) is relatively high, but a system that can be practically deployed will need an extremely high rejection rate of false alarms and provide very low missed detections. Fig. 1(c) presents the results using information fusion at feature level for radar plus accelerometer, i.e. by combining into a single feature vector the feature samples extracted from the radar and the accelerometer data. The overall classification accuracy across the ten activities is in this case 80.6%, and for most activities the classification accuracy has increased compared with the case of independent use of the radar sensor or the accelerometer sensor.

IV. RESULTS USING ADDITIONAL SENSORS

RGB-D sensors (e.g. Kinect) project infrared light into the space and detect the distortion of this pattern. Since these sensors produces large feature pools, PCA (Principle Component Analysis) that selects optimal combinations of the features with the largest possible variance is implemented to reduce the computational complexity in feature selection.

The results presented in Fig. 1(c) show persistent misclassifications for activities 3, 8, 9, 10, as well as for the pair

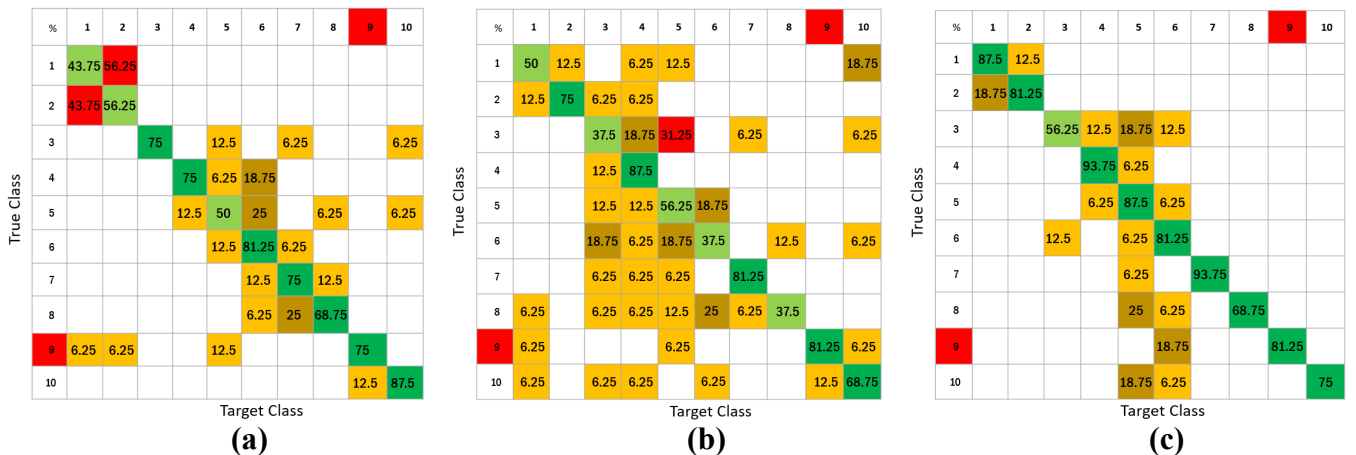


Fig. 1. Confusion matrices of radar data with SVM classifier (a), accelerometer with SVM classifier (b), and fusion radar plus accelerometer features with SVM classifier (c)

	%	1	2	3	4	5	6	7	8	9	10
1		87.5	6.25			6.25					
2			93.75								6.25
3		6.25		93.75							
4					100						
5					6.25	93.75					
6					6.25	12.5	75		6.25		
7					6.25		6.25	87.5			
8						18.75	6.25		75		
9										93.75	6.25
10											100

Fig. 2. Confusion matrix of sensor data fusion radar plus accelerometer and Kinect features with Quadratic-kernel SVM classifier

Accuracy with different sensor fusion and classifier

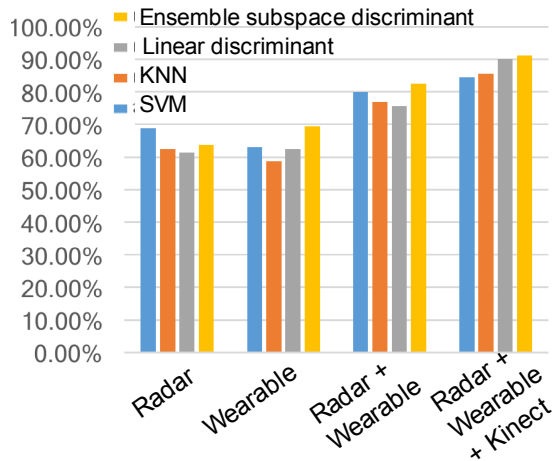


Fig. 3. Comparison of different sensors combination with different classifiers

of activities 1 and 2. The feature extracted from the Kinect sensor data are then added to those extracted from accelerometer and radar to investigate whether those results can be further improved. The resulting confusion matrix is shown in Fig. 2.

Fusing information from three sensors increases the classification accuracy to 86.9% with the quadratic-kernel SVM classifier, and up to 91.3% using an Ensemble classifier. The confusion matrix shown in Fig. 2 presents on average less misclassifications compared with the previous cases, although in some particular cases (e.g. activity 6) using a 3rd sensor decreases the accuracy compared with using only two, so optimal approaches to fuse information should be considered to avoid classifiers with low performance to decrease the overall performance. Fig. 3 summarises the average classification accuracy (assuming 16-fold cross-validation on the data from the 16 participants) for 4 different classifiers and for different combinations of sensors using feature-level information fusion. A clear trend of increasing accuracy when combining different sensors can be seen, and it is also interesting to observe the effect of using different types of classifier on data from the same sensor. Future work will build up on these preliminary results

and investigate in more detail the best approaches in terms of feature selection for data from each sensor and of information fusion.

V. CONCLUSION AND FUTURE WORK

Different research directions can build up on the preliminary results presented in this paper. For future work, additional data will be collected, involving 10 DOF (Degrees of Freedom) inertial measurement unit (three axes accelerometer, gyroscope, magnetic sensors and one pressure sensor or GPS) [12], more participants, more indoor scenarios, and more deployment geometries for the sensors. This includes different aspect angles of the Kinect and radar with respect to the participants' movements and trajectories, as well as different positions of the accelerometer sensor (e.g. on the wrist like for this paper, at the waist or chest or arms or thighs, or inside pockets) and multiple accelerometers. The integration of gyroscope and magnetic sensors together with the accelerometer and the other sensors will also be considered. On the signal processing side, different approaches to select features for each different sensor will be considered (for example using metrics such as entropy or Fisher's scores to rank all the possible features), as well as the effect of using different information fusion techniques on data from all the sensors (e.g. fusing at feature level, or at decision level taking into account the level of confidence of each separate classifier based on data from each individual sensor).

ACKNOWLEDGMENT

This work was partly supported by a STSM funded by the COST Action IC1303 AAEPEL (*Architectures, Algorithms and Platforms for Enhanced Living Environments*).

REFERENCES

- [1] Fabbri E, *et al.*, "New Tasks, Priorities, and Frontiers for Integrated Gerontological and Clinical Research", *Journal of the American Medical Directors Association*, vol. 16, no. 8, pp. 640-647, 2015.
- [2] Terroso, M., Rosa, N. and Torres Marques, "Physical consequences of falls in the elderly: a literature review from 1995 to 2010", *Eur Rev Aging Phys Act*, pp.11-51, 2014.
- [3] K. Chaccour, R. Darazi, A. H. El Hassani, and E. Andr s, "From Fall Detection to Fall Prevention: A Generic Classification of Fall-Related Systems," *IEEE Sensors Journal*, vol. 17, no. 3, pp. 812-822, 2017.
- [4] S. C. Mukhopadhyay, "Wearable Sensors for Human Activity Monitoring: A Review," *IEEE Sensors Journal*, vol. 15(3), 2015.
- [5] I. H. Lopez-Nava, M. M. Angelica, "Wearable Inertial Sensors for Human Motion Analysis: A review," *IEEE Sensors Journal*, vol.16 (22), 2016.
- [6] S. Wen, *et al.*, "A Wearable Fabric-Based RFID Skin Temperature Monitoring Patch," in *Proc. of IEEE Sensors Conference*, pp. 1-3, 2016.
- [7] F. Erden, *et al.*, "Sensors in Assisted Living: A survey of signal and image processing methods," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 36-44, 2016.
- [8] E. Cippitelli, F.Fioranelli, E. Gambi, S. Spinsante, "Radar and RGB-Depth sensors for fall detection: a review", *IEEE Sensors Journal*, vol. 17, no. 12, pp. 3585-3604, 2017.
- [9] D. Figo *et al.* "Preprocessing techniques for context recognition from accelerometer data", *Personal Ubiquitous Computing*, vol. 14, no.7, 2010.
- [10] Lee, M-H., *et al.* "Physical Activity Recognition Using a Single Tri-Axis Accelerometer", *Proc. World Congress Eng. & Computer Science*, 2009.
- [11] E.M.Tapia, "Using Machine Learning for Real-time Activity Recognition and Estimation of Energy Expenditure", MIT PhD Thesis, 2008.
- [12] H. Heidari, *et al.*, "CMOS vertical Hall magnetic sensors on flexible substrate," *IEEE Sensors Journal*, vol. 16, pp. 8736-8743, 2016.